

REMARKS/ARGUMENTS

Claims 19-36 stand in the present application, claims 19, 23, 27, 29 and 33 having been added. Reconsideration and favorable action is respectfully requested in view of the above amendments and the following remarks.

In the Office Action, the Examiner has objected to the title of the invention. As noted above, Applicants have amended the title of the invention to correct the deficiency pointed out by the Examiner.

The Examiner has rejected claims 23 and 33 under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point and distinctly claim the subject matter which Applicants regard as the invention. As noted above, Applicants have amended claims 23 and 33 in the manner suggested by the Examiner in order to correct this deficiency.

The Examiner has rejected claims 19 and 27 under 35 U.S.C. § 101 for being non-statutory. Applicants respectfully traverse the Examiner's § 101 rejection of the claims. It is respectfully submitted that claim 19 clearly transforms data to a different state or thing when it receives data, analyzes the data, buffers the received audio stream and performs speech recognition process in order to determine an intended receiver of the audio stream. Clearly each of these steps or operations transform data and thus meets the statutory definition. In any event, Applicants have amended claim 19 in order to recite a push-to-talk communications method for use with a push-to-talk communications device. Accordingly, it is respectfully submitted that amended claim 19 now more clearly complies with 35 U.S.C. § 101. With respect to claim 27, Applicants

have amended the claim to more clearly recite that a tangible computer medium containing a computer program is arranged in order to perform the steps of method claim 19. Accordingly, claim 27 as amended is also believed to comply with 35 U.S.C. § 101.

The Examiner has rejected claims 19 and 27-29 under 35 U.S.C. § 103(a) as being unpatentable over Vysotsky et al. ("Vysotsky") in view of Dailey, has rejected claims 20-24 and 30-34 under 35 U.S.C. § 103(a) as being unpatentably over Vysotsky in view of Dailey and further in view of Guilhufe et al. ("Guilhufe"), has rejected claims 26 and 36 under 35 U.S.C. § 103(a) as being unpatentable over Vysotsky in view of Dailey and further in view of Schrage, and has rejected claims 25 and 35 under 35 U.S.C. § 103(a) as being unpatentable over Vysotsky in view of Dailey and further in view of Guilhufe and further in view of Salazar et al. ("Salazar"). Applicants respectfully traverse the Examiner's § 103 rejections of the claims.

In answer to the Office Action's response to Applicants' previous arguments, Applicants respectfully submit the following remarks. The Examiner alleges that the "push-to-talk" feature of Applicants' invention is not recited in the present claims. See, Office Action at page 2. To the contrary, Applicants submit that the push-to-talk feature is recited in both independent claims 19 and 29. For example, claim 19 requires "receiving at a router server in a communications network an audio stream containing an utterance **when a user presses a button on a user device and starts to talk.**" Similar language appears in claim 29. Although the claim language clearly shows that claims 19 and 29 require a push-to-talk method and system, Applicants have amended

claim 19 to recite a "push-to-talk communications method" and has similarly amended claim 29 to recite a "push-to talk communications system."

The Examiner continues to allege that Geilhufe teaches use of a half-duplex system and cites as support for this allegation "column 22, lines 57-65 where a 'random delay of up to 2 seconds relative silence' is described." *Id.* However, Geilhufe itself contradicts the Examiner's allegation. For example, Geilhufe states at column 19, lines 18 – 20 that: *"The <silence> is a period of **relative** silence during which the user is not speaking although background noise and background speech may still be present."* Geilhufe also states at column 19, lines 38 – 40 that: *"Requiring <silence> before detection of <name> helps prevent false detection of <name> during normal conversational speech"*. Geilhufe further states at column 22, line 66 – column 23, line 23 that *"during the random delay of up to 2 seconds, each responding voice controlled device **listens** for another voice controlled device's response . . . "* Hence, in Geilhufe, background noise and speech can be heard during this delay period and the devices are also able to listen for responses from other devices during this delay and it is, therefore, absolutely clear from Geilhufe's disclosure that the random delay is not caused by a shift from "transmission for input" to a "transmission for response" but is introduced for avoiding false detection of a name during normal conversation or for avoiding conflicts between different devices.

The Office Action also incorrectly states that a WLAN is a half duplex packet based system and since the system in Geilhufe can be connected to the Internet this also implies that the system in Geilhufe is half duplex. Applicants respectfully disagree

for the follow reasons. WLAN may have a half duplex transport layer, but that is irrelevant. The half duplex nature that the Office Action is talking about is at completely the wrong layer in the IP stack. Applicants' invention is concerned with a half duplex voice interaction which means that both people cannot talk at the same time. Most VoIP telephone calls will pass over some transport layer connection that is half duplex at some point, but they are still full duplex voice calls – both people can speak (and be heard) at the same time. Hence, the arguments and assertions that Geilhufe discloses a half-duplex system are not correct.

Regarding the prior argument concerning "group created on the fly," Applicants respectfully submit this is not a feature required by the present claims nor on which arguments for patentability are predicated. From Applicants' response to the first Office Action it is clear that the explanation that the group is created on the fly only relates to former claims 4 and 14, now claims 22 and 32, which claim an embodiment of the invention related to group calls, and that the phrase was used to explain to the Examiner how the invention works. The features that enables the group call, namely performing a speech recognition process on the received audio stream to recognize the utterance contained therein; determining, if possible, an intended receiver of the audio stream in dependence on the recognized utterance; and if an intended receiver was determined, transmitting the audio stream to the determined intended receiver, are stated in claim 19. Other embodiments are not related to group calls; it is clear from the application as filed that the caller can use the invention for calling a single person or a group of persons. In both cases the intended receiver(s) are determined based on the

features of claim 19. Hence, Applicants disagree with the Examiner that the phrase “group created on the fly” should or need be incorporated into the claims.

Turning to the prior art rejections, Vysotsky describes a fairly standard voice dialling application. The method and system permit a customer to place a call by speaking a person’s name which serves as a destination identifier without having to speak an additional command or steering word to place the call. Vysotsky uses two parallel speech recognition systems, the first being speaker independent for recognising commands such as “cancel call forwarding” and the second being speaker dependent for recognising names in the received message, e.g. to find who to call.

Contrary to Applicants’ invention, Vysotsky deals with duplex telephone systems and methods and a call always starts with a separate dialling phase. For example, a confirmation message confirming which person the system will call or what command he/she wants to perform is always played back to a user before the initiation of a command or call. See, Vysotsky at column 10, lines 48 – 59 and column 11, lines 30 – 34. Hence, in Vysotsky there is a separate dialling phase when a caller first says “John” which initiates a search in the speaker dependent storage for the name John and an associated number and after the number is found a call is initiated. What Vysotsky aims to achieve is to get rid of predefined commands or steering words when initiating a call since these places constraints on a user in terms of *how* a user must format a spoken request to dial a call.

Dailey discloses (i) a single button conference call initiator (to a predefined set of callees) or (ii) a single button initiator for a half duplex group audio delivery system

where the group of recipients is predefined. Dailey discloses a Push-to-talk system but a user has to pre-define which users belong to a group. A group call is broadcast in several cells simultaneously on a common channel, and there is no requirement that the system page the terminals of the group. This can reduce the overhead involved in setting up a group call, as a group call can be setup by simply receiving a predetermined type of call origination message from a terminal of a group, and responsively transmitting a group call traffic channel designation message to the other terminals of the group. This can be achieved by transmitting the group call traffic channel designation message over a paging sub channel mapped onto the common control channel (DCCH) that is broadcast by each radio head of the system. Terminals receiving a group call traffic channel designation message can then simply tune to the traffic channel designated in the message, e.g., to the frequency/slot identified in the message, and await reception of voice data over the common traffic channel. Preferably, no ring will occur at these terminals; they simply will unmute their speakers and begin generating audio signals upon receipt of voice data over the common traffic channel. Hence, Dailey does not disclose a determination step determining one or a plurality of intended receivers.

In Applicants' invention a call to one or a number of different people is created when a user presses a button on his device and says "Hi John, I heard you are . . ." or "Hi Peter, Simon and Roger, I'd like to discuss . . ." Hence, there is no separate dialling phase in Applicants' invention, but the phrase a user is saying is used for determining

an intended receiver (John) or receivers (Peter, Simon, Roger) and then this same phrase is routed to the intended receiver or receivers.

As stated above, Vysotsky implements a separate dialling phase where a caller first has to say "John," then wait for the system to find the number to John and then initiate the call to John. In Dailey the "dialling phase" consists of the broadcasting of a message that a group call for a predefined group will take place on a certain traffic channel. A member of the group that hears the broadcast then tunes to the designated traffic channel. Determining intended receivers of a call and then routing a call to them is thus not implemented in Dailey.

Hence, a skilled person in the art implementing the teachings of Dailey into Vysotsky in order to enable Push-to-talk calls in Vysotsky, would end up with a system for group calls where the initial dialling phase using speech recognition is discarded and instead replaced with a system that broadcasts a message on a traffic channel to a predefined group of people. Accordingly, the skilled person in the art would be directed away from Applicants' invention by combining Vysotsky and Dailey and would definitely not end up with a method or system for *"performing a speech recognition process on a received audio stream to recognize an utterance contained therein; determining, if possible, an intended receiver of the audio stream in dependence on the recognized utterance; and if an intended receiver was determined, transmitting said audio stream to the determined intended receiver using a half-duplex communications service provided by a packet-switched network,"* as required by present claim 19. Accordingly, claim 19

and its respective dependent claims patentably define over Vysotsky and Dailey taken singly or in combination.

With respect to present system claim 29, the Examiner states that claim 10 in Vysotsky teaches a device that corresponds to the claimed invention.

A device for responding to speech, comprising:

means for performing speaker independent speech recognition on the speech to detect the presence of a spoken command in the speech;

means for performing speaker dependent speech recognition on the speech to detect a non-command word in the speech, the speaker independent and speaker dependent speech recognition means operating in parallel;
and

a device for performing an action in response to the detection of a command by the speaker independent speech recognition means.

See, Claim 10 of Vysotsky (emphasis supplied). In Applicants' system there are not two parallel speech recognition means for determining both speaker independent and speaker dependent speech. Further, Applicants' system comprises *a router server in a network*; which the device in Vysotsky does not; and lastly the device in Vysotsky is arranged to perform an action in the response to the detection of *a command* by one of the speech recognition means; however, since no commands are ever uttered in Applicants' invention this part of the Vysotsky device is definitely not included in the present invention.

Accordingly, claim 29 and its respective dependent claims also patentably define over the cited art taken singly or in combination. Applicants further submit that since the

present independent claims 19 and 29 patentably define over Vysotsky and Dailey, the secondary references need not be addressed at this time. However, some additional comments concerning claim rejections further based on the Geilhufe reference are discussed below.

Furthermore, with regard to claim 20, the Examiner states that Geilhufe teaches "indicating the one or more possible intended receivers to a user." See, Office Action at page 7. However, in Geilhufe, **all** the appliances receive the audio, so the Examiner's use of the phrase "intended receiver" is misleading; "intended actioner" would be more accurate, e.g., there is no receiver specific audio routing in Geilhufe. Hence, combining the teachings of Vysotsky, Dailey and Geilhufe would not have lead to present claim 20.

With regard to claims 22 and 32 it has already been explained that Dailey teaches broadcasting of a group call over a communications channel to a predefined group of receivers and hence there is no determining step that determines a plurality of intended receivers either in Dailey or Vysotsky. The Examiner further relies on Geilhufe, asserting that "it would have been obvious to someone in the art . . . to modify Dailey and Vysotsky with the Geilhufe device to provide selection and transmission to multiple receivers through speech selection." However, as stated above **all devices** in Geilhufe receive a command from a user although only a few of the devices will react by performing a task in accordance with the command; hence, there is no selection of devices followed by transmission to only the selected devices in Geilhufe.

With regard to claims 23 and 33, the Examiner asserts that the syntax used in Geilhufe implies that the speech recognition is only performed on a portion of the

received audio stream. The assertion that Geilhufe suggests that only the early part of the utterance has to be recognised applies a post-facto-analysis. Geilhufe says no such thing. Just because the appliance name is recognized in the early part of utterance doesn't mean that the system does not have to recognize the whole utterance, or indeed even know the result of the early part before finishing recognizing the whole thing. In fact, with the grammar specified in Geilhufe, the system would have no choice but to recognize the entire utterance before being able to determine the appliance name.

A key advantage of Applicants' invention is that there is no separate dialling phase, but a user's phrase (utterance) is used for recognizing an intended receiver or receivers and then this same phrase is routed to this recognized intended receiver or receivers, which results in quicker call setups. None of the cited references taken singly or in any combination enable this feature.

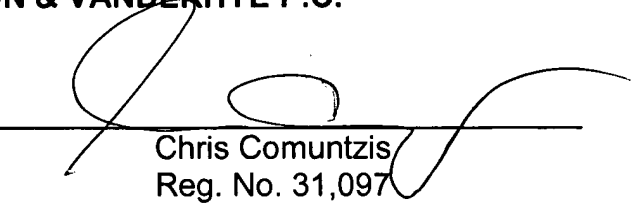
Therefore, in view of the above amendments and remarks, it is respectfully requested that the application be reconsidered and that all of claims 19-36, standing in the application be allowed and that the case be passed to issue. If there are any other issues remaining which the Examiner believes could be resolved through either a supplemental response or an Examiner's amendment, the Examiner is respectfully requested to contact the undersigned at the local telephone exchange indicated below.

RINGLAND et al.
Appl. No. 10/581,290
April 6, 2009

Respectfully submitted,

NIXON & VANDERHYE P.C.

By: _____


Chris Comuntzis
Reg. No. 31,097

CC:lmr
901 North Glebe Road, 11th Floor
Arlington, VA 22203-1808
Telephone: (703) 816-4000
Facsimile: (703) 816-4100